



ARTIGO ORIGINAL

Sistema inteligente de apoio à decisão tática no futebol: engenharia de dados e *dashboards* interativos aplicados ao campeonato brasileiro

José Lucas Vasconcelos de Lucena¹; Patricia Medyna Lauritzen de Lucena Drumond²; Carlos Futino Barreto³

Como Citar:

DE LUCENA, José Lucas Vasconcelos; DRUMOND, Patricia Medyna Lauritzen de Lucena; BARRETO, Carlos Futino. Sistema inteligente de apoio à decisão tática no futebol: engenharia de dados e dashboards interativos aplicados ao campeonato brasileiro. Revista Sociedade Científica, vol. 8, n. 1, p. 2372-2397, 2025. <https://doi.org/10.61411/rsc2025117818>

DOI: 10.61411/rsc2025117818

Área do conhecimento:

Engenharias; Ciências da Computação

Sub-área:

Engenharia de Software; Engenharia de Dados

Palavras-chaves: Inteligência Artificial. Futebol. Análise Tática. Aprendizado de Máquina. Automação Esportiva.

Publicado: 19 de novembro de 2025

Resumo

O avanço da inteligência artificial e da ciência de dados tem transformado o futebol moderno, especialmente na análise tática e na tomada de decisão em tempo real. Este estudo apresenta o desenvolvimento do Tático Pro, um sistema inteligente que integra PostgreSQL, FastAPI, LangChain e LlamaIndex para gerar relatórios automatizados e recomendações baseadas em aprendizado de máquina. O objetivo é avaliar a aplicabilidade da IA como ferramenta de apoio técnico na identificação de padrões de jogo e vulnerabilidades adversárias. A pesquisa adota abordagem exploratória e aplicada, fundamentada em revisão bibliográfica e experimentação prática com dados do Campeonato Brasileiro de 2024, processados por consultas dinâmicas e dashboards interativos. O sistema traduziu dados complexos em insights táticos acessíveis, permitindo consultas em linguagem natural e respostas contextualizadas. Os resultados indicam que a integração entre análise de dados e IA eleva a precisão das decisões táticas, otimizando o tempo de análise e fortalecendo a gestão estratégica no futebol brasileiro.

Intelligent Decision Support System in Soccer: Data Engineering and Interactive Dashboards Applied to the Brazilian Championship

Abstract

¹Instituto de Ensino Superior - ICEV, Teresina-PI, Brasil. Email: jlucena@icev.edu.br

²Universidade Federal do Piauí - UFPI, Teresina-PI, Brasil. Email: patricia@ufpi.edu.br

³Instituto de Ensino Superior - ICEV, Teresina-PI, Brasil. Email: cfutino@icev.edu.br



The advancement of artificial intelligence and data science has transformed modern soccer, particularly in tactical analysis and real-time decision-making. This study presents the development of Tático Pro, an intelligent system that integrates PostgreSQL, FastAPI, LangChain, and LlamaIndex to generate automated reports and machine learning–based recommendations. The objective is to evaluate the applicability of AI as a support tool for technical teams in identifying game patterns and opponent vulnerabilities. The research adopts an exploratory and applied approach, supported by a literature review and practical experimentation using data from the 2024 Brazilian Championship, processed through dynamic queries and interactive dashboards. The system translated complex datasets into accessible tactical insights, enabling natural language queries and contextualized responses. The results indicate that integrating data analysis and AI enhances the precision of tactical decisions, optimizes analysis time, and strengthens strategic management in Brazilian soccer.

Keywords: Artificial Intelligence. Soccer. Tactical Analysis. Machine Learning. Sports Automation.

1. Introdução

O futebol consolidou-se como o esporte mais popular do Brasil e, ao mesmo tempo, como um espaço privilegiado para a aplicação de tecnologias emergentes na análise de desempenho. A crescente disponibilidade de dados oriundos de sistemas de rastreamento, sensores, plataformas digitais e registros estatísticos abriu caminho para o desenvolvimento de sistemas inteligentes capazes de transformar dados brutos em informações estratégicas para treinadores e analistas. Essa evolução acompanha uma tendência global do esporte de alto rendimento, no qual a integração de engenharia de dados, inteligência artificial e visualização interativa possibilita análises cada vez mais profundas sobre padrões táticos e técnicos. Trabalhos recentes demonstram que o uso de algoritmos de aprendizado de máquina e modelagem estatística pode prever posições finais em campeonatos, identificar fatores críticos para vitórias e apontar indicadores de



performance coletiva, revelando que a ciência de dados vem se tornando essencial na tomada de decisão esportiva [2.,4.].

Entretanto, as plataformas atualmente disponíveis no mercado, como Wyscout e SofaScore, concentram-se em fornecer estatísticas descritivas, sendo limitadas no que tange à geração de insights preditivos e estratégicos. Além disso, apresentam custos elevados e barreiras de acesso que dificultam sua utilização por clubes brasileiros de médio e pequeno porte, o que gera desigualdade competitiva e limita a democratização do uso da análise avançada de dados no futebol nacional. Pesquisas como a de Lolli *et al.* [3.] evidenciam que, embora a indústria do futebol já esteja fortemente orientada pela ciência de dados, ainda existem lacunas significativas na integração de diferentes camadas informacionais — como dados individuais, setoriais e coletivos — em plataformas acessíveis e de fácil interpretação. Essa ausência reforça a necessidade de sistemas adaptados à realidade brasileira, que considerem tanto a escassez de recursos de determinados clubes quanto a especificidade competitiva do Campeonato Brasileiro, um dos mais equilibrados do mundo.

Nesse contexto, estudos de análise visual e dashboards interativos emergem como uma alternativa promissora para ampliar a compreensão das dinâmicas do jogo e otimizar estratégias táticas. Santos *et al.* [6.] destacam que ferramentas baseadas em visualização interativa permitem identificar comportamentos coletivos e individuais, revelando tendências que não são facilmente percebidas em análises convencionais. Da mesma forma, Evers *et al.* [1.] demonstram que técnicas de visual analytics são eficazes para avaliar desempenho de jogadores de forma intuitiva, aproximando os resultados da análise científica das necessidades práticas de treinadores e comissões técnicas. Essa convergência entre big data analytics e ferramentas visuais fundamenta a proposta deste estudo, que visa desenvolver uma plataforma capaz de apoiar decisões táticas a partir de dados estruturados e apresentados em dashboards acessíveis.

O desafio, portanto, não está apenas em coletar grandes volumes de dados, mas em garantir sua padronização, consistência e integração para que possam gerar conhecimento útil em tempo real. Mănescu *et al.* [5.] ressaltam que o valor da análise



em esportes de elite depende da construção de pipelines robustos de ingestão, transformação e orquestração de dados, de modo que informações complexas possam ser entregues em tempo hábil para subsidiar a tomada de decisão. Nesse sentido, a engenharia de dados assume papel central, pois organiza as camadas de dados em modelos relacionais, assegura qualidade informacional e habilita a criação de indicadores confiáveis. Associado a isso, o uso de agentes inteligentes baseados em linguagem natural amplia o alcance do sistema, ao possibilitar que técnicos e analistas obtenham respostas diretas a questões complexas, sem a necessidade de conhecimento especializado em estatística ou programação [6.,7.].

Diante desse cenário, o presente artigo propõe a construção de um sistema inteligente de apoio à decisão tática, aplicado especificamente ao Campeonato Brasileiro de 2024, que integra técnicas de engenharia de dados, visualização interativa e inteligência artificial. O estudo se organiza como um estudo de caso, estruturado em etapas que incluem: preparação e normalização de um banco de dados relacional, implementação de pipelines de ingestão e transformação, desenvolvimento de dashboards interativos e criação de um protótipo de agente inteligente para consultas em linguagem natural. A proposta busca responder à seguinte questão central: como transformar dados brutos do futebol em informações estratégicas, acessíveis e úteis para técnicos brasileiros? Ao longo do trabalho, serão discutidos os fundamentos teóricos, o processo de desenvolvimento da plataforma e a validação do sistema em cenários simulados, contribuindo para a inovação tecnológica no futebol brasileiro e para a democratização da análise avançada de dados no esporte.

2. Referencial teórico

A incorporação de tecnologias digitais e inteligência artificial ao esporte tem transformado o modo como o desempenho atlético é compreendido e analisado. No futebol, essa transformação é impulsionada pela consolidação da engenharia de dados como ferramenta estruturante para a coleta, processamento e análise de grandes volumes de informações. Segundo Mănescu *et al.* [5.], a engenharia de dados é a base da análise



esportiva moderna, pois garante a integridade e a escalabilidade dos sistemas que sustentam a inteligência aplicada ao jogo. Essa abordagem permite que clubes e analistas convertam dados dispersos em indicadores úteis para a formulação de estratégias táticas.

A construção de sistemas inteligentes no futebol requer pipelines capazes de integrar dados de diferentes origens, como sensores, câmeras, registros manuais e APIs oficiais. Lolli *et al.* [3.] apontam que a automação desses fluxos de dados é fundamental para reduzir erros humanos e aumentar a confiabilidade das análises. O uso de ferramentas como Airbyte, dbt e Airflow exemplifica o avanço na orquestração de dados esportivos, permitindo que todo o processo de ingestão e transformação seja realizado de forma contínua e auditável, gerando modelos relacionais consistentes.

O aprendizado de máquina é um dos principais aliados da ciência de dados esportiva. Kleina *et al.* [2.] demonstraram que algoritmos supervisionados, como redes neurais e máquinas de vetores de suporte, são capazes de prever a classificação final dos times no Campeonato Brasileiro com base em dados das rodadas iniciais. Esses modelos preditivos abrem espaço para análises mais estratégicas e embasadas, substituindo a intuição por inferências baseadas em evidências.

Os modelos de aprendizado supervisionado também têm sido aplicados à análise do comportamento coletivo e individual dos atletas. Santos *et al.* [6.] mostraram que a análise de dados posicionais, combinada à visualização interativa, possibilita identificar padrões táticos e falhas estruturais que escapam à observação visual humana. Essa integração entre estatística e visualização amplia o potencial analítico das comissões técnicas.

A visualização de dados, quando aliada a dashboards interativos, torna-se uma poderosa ferramenta cognitiva para os profissionais do esporte. Evers *et al.* [1.] destacam que interfaces gráficas dinâmicas favorecem o raciocínio visual e melhoram a comunicação entre técnicos, analistas e atletas. A partir dessas visualizações, é possível compreender relações espaciais e temporais complexas, transformando o processo de análise em uma atividade colaborativa e acessível.



A complexidade do futebol brasileiro exige soluções adaptadas ao contexto nacional. Maia *et al.* [4.] afirmam que o Campeonato Brasileiro é caracterizado por equilíbrio técnico, variação climática e intensa sequência de jogos, o que torna a modelagem estatística mais desafiadora. Modelos estocásticos permitem representar incertezas e simular cenários que incorporam variáveis como mando de campo, número de faltas, cartões e desgaste físico.

A qualidade dos dados é outro pilar essencial para a confiabilidade das análises. Mănescu *et al.* [5.] enfatizam que falhas na padronização ou inconsistências em atributos podem distorcer resultados preditivos e gerar conclusões equivocadas. A normalização de dados e o tratamento de valores ausentes são práticas fundamentais para evitar vieses e assegurar comparabilidade entre diferentes jogos e temporadas.

A integração de múltiplas fontes de dados eleva o potencial das análises esportivas. Lolli *et al.* [3.] observam que a combinação de informações táticas, fisiológicas e médicas pode revelar relações inéditas entre carga física, posicionamento e desempenho. Essa visão holística fortalece a ideia de um ecossistema digital do futebol, no qual a informação circula de forma integrada entre departamentos.

Sistemas inteligentes de apoio à decisão combinam fundamentos de estatística, aprendizado de máquina e design de interfaces para fornecer recomendações estratégicas. Xu, Zhang e Wang [8.] destacam que plataformas baseadas em análise de dados podem orientar treinadores sobre formações mais eficientes ou ajustes táticos conforme o desempenho do adversário. Essa abordagem coloca a inteligência artificial como coadjuvante direto no processo decisório esportivo.

A personalização dos dashboards também é um fator determinante para sua eficácia. Evers *et al.* [1.] demonstram que interfaces adaptáveis às necessidades de diferentes usuários tornam o processo analítico mais eficiente. Técnicos demandam painéis sintéticos e comparativos, enquanto analistas preferem gráficos detalhados e históricos completos. Essa diferenciação garante maior engajamento e utilidade prática.



A tomada de decisão em tempo real é um dos maiores desafios enfrentados pelos analistas de futebol. Mănescu *et al.* [5.] ressaltam que pipelines automatizados permitem identificar mudanças de desempenho durante os jogos e emitir alertas instantâneos. Essa funcionalidade transforma o processo de análise em uma ferramenta ativa, capaz de influenciar intervenções imediatas no campo.

O uso de agentes conversacionais baseados em linguagem natural vem ganhando destaque nas plataformas de análise esportiva. Silva [7.] descreve que a integração de modelos de NLP em sistemas de apoio à decisão facilita consultas complexas, permitindo que técnicos perguntem em linguagem comum sobre estatísticas ou padrões de jogo. Essa tecnologia reduz barreiras técnicas e aumenta a acessibilidade do sistema.

A validação de sistemas inteligentes requer experimentação prática. Kleina *et al.* [2.] validaram seus modelos preditivos com dados reais do Campeonato Brasileiro, obtendo resultados superiores a 80% de acurácia. Essa metodologia reforça o valor da pesquisa aplicada e demonstra que soluções acadêmicas podem ser adaptadas à realidade dos clubes brasileiros.

Os modelos estatísticos mais recentes utilizam combinações de distribuições de Poisson e cadeias de Markov para prever probabilidades de resultados. Maia *et al.* [4.] argumentam que tais modelos são capazes de representar a natureza probabilística do futebol, considerando fatores como vantagem do mandante e desempenho ofensivo. Essas abordagens se aproximam das técnicas utilizadas por grandes clubes europeus.

O conceito de big data esportivo amplia a escala da análise, permitindo o processamento de milhões de registros em tempo reduzido. Lolli *et al.* [3.] afirmam que clubes de elite utilizam arquiteturas distribuídas em nuvem e sistemas de data lakes para consolidar dados provenientes de sensores e vídeo. No contexto brasileiro, a adoção de soluções de código aberto representa uma alternativa viável para reduzir custos e democratizar o acesso à tecnologia.

A explicabilidade dos algoritmos é um aspecto cada vez mais valorizado. Xu, Zhang e Wang [8.] explicam que a inteligência artificial explicável (XAI) torna as decisões dos modelos mais transparentes, permitindo que treinadores compreendam os



fatores que levaram a determinada recomendação. Isso aumenta a confiança na tecnologia e favorece sua adoção.

A visualização interativa também desempenha um papel educativo no ambiente esportivo. Santos *et al.* [6.] afirmam que dashboards visuais promovem aprendizado contínuo e favorecem o desenvolvimento tático dos atletas. Essa interação visual auxilia no entendimento coletivo do jogo e na comunicação entre os membros da equipe técnica.

No campo da engenharia de software aplicada ao esporte, a modularidade é considerada uma boa prática essencial. Mănescu *et al.* [5.] observam que plataformas modulares permitem incorporar novos indicadores e dados sem a necessidade de reformulação total da arquitetura. Essa flexibilidade garante longevidade e adaptabilidade ao sistema.

O desenvolvimento de sistemas web esportivos requer atenção à experiência do usuário. Evers *et al.* [1.] apontam que dashboards com usabilidade aprimorada incentivam a exploração de dados e reduzem a curva de aprendizado. Ferramentas modernas de front-end, como Next.js e bibliotecas de visualização em JavaScript, permitem construir interfaces dinâmicas e responsivas.

A democratização do acesso à análise de dados é um tema recorrente na literatura. Lolli *et al.* [3.] afirmam que a adoção de soluções abertas e gratuitas pode reduzir a desigualdade tecnológica entre clubes, tornando a análise tática acessível a instituições de menor orçamento. Essa visão social da tecnologia reforça a importância da pesquisa aplicada no contexto esportivo nacional.

Os agentes inteligentes também contribuem para a eficiência operacional dos departamentos de análise. Silva [7.] descreve que chatbots esportivos podem responder rapidamente a dúvidas sobre estatísticas de jogadores e tendências táticas, otimizando o tempo da comissão técnica. Essa integração melhora o fluxo de comunicação e aumenta a produtividade.



O futuro da análise esportiva aponta para a convergência entre inteligência artificial generativa, visualização preditiva e processamento em tempo real. Xu, Zhang e Wang [8.] preveem que sistemas híbridos, capazes de aprender continuamente com novos dados, redefinirão o conceito de futebol inteligente, aproximando ciência de dados e prática esportiva.

A validação empírica das ferramentas desenvolvidas é indispensável para garantir credibilidade científica. Mănescu *et al.* [5.] afirmam que experimentos em competições reais são o método mais eficaz para avaliar a robustez de sistemas de apoio à decisão. O uso do Campeonato Brasileiro como estudo de caso fortalece a aplicabilidade prática deste tipo de pesquisa.

A literatura demonstra que a combinação entre engenharia de dados, inteligência artificial e visualização interativa representa uma evolução significativa na forma de compreender o futebol. Ao unir tecnologia e conhecimento tático, é possível construir sistemas que ampliam a capacidade de análise e tornam o processo decisório mais objetivo e fundamentado.

A literatura especializada também tem explorado o conceito de integração entre análise de desempenho e inteligência situacional, que consiste na capacidade do sistema de compreender o contexto do jogo e oferecer recomendações adaptativas. Xu, Zhang e Wang [8.] explicam que esse tipo de inteligência permite que o sistema interprete variações táticas em tempo real e ajuste suas previsões conforme a evolução do jogo, o que representa um avanço em relação às análises estáticas tradicionais. Essa abordagem pode ser determinante para a identificação de padrões contextuais, como a reação de uma equipe após sofrer um gol ou mudanças de postura tática diante de adversários específicos.

O uso de dashboards preditivos vem se consolidando como um diferencial estratégico para equipes técnicas que buscam decisões mais rápidas e embasadas. Santos *et al.* [6.] descrevem que visualizações dinâmicas baseadas em dados de eventos, como finalizações, interceptações e passes decisivos, tornam a análise mais intuitiva e acessível. A representação gráfica dos indicadores permite que técnicos e analistas



identifiquem imediatamente tendências de desempenho, fragilidades defensivas e oportunidades ofensivas, reduzindo o tempo necessário para tomadas de decisão.

A interdisciplinaridade é outro aspecto central abordado na literatura sobre engenharia de dados aplicada ao futebol. Mănescu *et al.* [5.] defendem que a colaboração entre engenheiros de software, estatísticos, cientistas de dados e profissionais de educação física é indispensável para o sucesso de sistemas analíticos esportivos. Essa integração de saberes favorece o desenvolvimento de soluções completas, que contemplam tanto a precisão técnica das análises quanto a aplicabilidade prática no campo de jogo.

Os avanços tecnológicos também têm impulsionado o uso de modelos híbridos, que combinam aprendizado supervisionado, redes neurais profundas e modelos de simulação probabilística. Maia *et al.* [4.] relatam que essa integração oferece previsões mais robustas sobre o resultado de partidas e o comportamento coletivo das equipes. Esses modelos são capazes de identificar interdependências complexas entre variáveis como posse de bola, intensidade de pressão e tempo de recuperação, gerando recomendações de alto valor estratégico.

A automação analítica é uma tendência consolidada no ambiente esportivo internacional. Lolli *et al.* [3.] apontam que clubes de elite europeus utilizam pipelines completamente automatizados, nos quais a coleta, transformação e visualização dos dados ocorrem de forma integrada e contínua. Essa abordagem minimiza falhas operacionais e garante que as informações estejam sempre atualizadas, permitindo que decisões táticas sejam tomadas com base em dados recentes e confiáveis.

Além da automação, a literatura destaca a relevância da interpretação contextual dos resultados. Segundo Evers *et al.* [1.], dashboards sem contexto podem gerar conclusões equivocadas, já que métricas isoladas não capturam a complexidade do jogo. A correlação entre dados estatísticos e a leitura subjetiva do técnico é, portanto, indispensável para que as recomendações geradas por sistemas inteligentes tenham aplicabilidade real.



Outra contribuição recente é o desenvolvimento de modelos de aprendizado por reforço aplicados à análise tática. Kleina *et al.* [2.] observam que esses algoritmos são capazes de aprender por meio da repetição de simulações, aprimorando-se progressivamente para sugerir estratégias mais eficientes. Essa metodologia pode ser utilizada, por exemplo, para simular o impacto de substituições ou alterações de formação, oferecendo uma base empírica para decisões tradicionalmente tomadas de forma intuitiva.

A análise espacial dos movimentos em campo também tem sido amplamente explorada. Santos *et al.* [6.] descrevem que a análise baseada em coordenadas espaciais dos jogadores possibilita compreender a ocupação de espaços, a compactação das linhas e as transições entre defesa e ataque. Esses dados, quando visualizados em painéis interativos, permitem que técnicos identifiquem zonas de vulnerabilidade e ajustem o posicionamento da equipe em treinos e partidas.

A engenharia de dados esportiva tem evoluído para além da coleta e modelagem, incorporando princípios de segurança e governança da informação. Mănescu *et al.* [5.] enfatizam que a conformidade com políticas de proteção de dados é fundamental para garantir a integridade das bases utilizadas nas análises. Esse cuidado é especialmente importante em sistemas que lidam com dados sensíveis de atletas, como informações médicas e fisiológicas.

Pesquisas contemporâneas indicam que a análise de sentimento e a mineração de texto também estão sendo incorporadas aos sistemas de apoio à decisão tática. Xu, Zhang e Wang [8.] relatam que o uso de processamento de linguagem natural permite analisar entrevistas, comentários e redes sociais para identificar o impacto psicológico de eventos esportivos sobre o desempenho dos jogadores. Essa perspectiva amplia o escopo da análise, agregando dimensões emocionais e comportamentais à modelagem estatística.

A literatura recente converge para a ideia de que o futuro dos sistemas inteligentes de análise esportiva depende da integração total entre ciência de dados, inteligência artificial e interfaces acessíveis. Lolli *et al.* [3.] defendem que essa



integração será responsável por transformar o futebol em um laboratório de inovação tecnológica contínua, no qual cada jogo se tornará uma fonte viva de aprendizado. Nesse contexto, o sistema proposto neste estudo posiciona-se como uma contribuição relevante ao aplicar conceitos de engenharia de dados e visualização interativa especificamente ao contexto do Campeonato Brasileiro, fortalecendo o vínculo entre pesquisa acadêmica e prática esportiva nacional.

A revisão de literatura confirma que a ciência de dados aplicada ao futebol brasileiro tem potencial para democratizar o acesso à informação, otimizar estratégias táticas e fortalecer a cultura analítica nos clubes. O desenvolvimento de um sistema inteligente de apoio à decisão, como o proposto neste estudo, representa um passo relevante na integração entre inovação tecnológica e gestão esportiva.

3. **Metodologia**

A metodologia deste estudo fundamenta-se em uma abordagem exploratória e aplicada, cujo objetivo é desenvolver, validar e analisar o desempenho do sistema Tático Pro, voltado ao apoio à decisão tática no futebol por meio da integração entre engenharia de dados, visualização interativa e inteligência artificial. O projeto adota um estudo de caso aplicado ao Campeonato Brasileiro de 2024, selecionado por sua alta complexidade competitiva e disponibilidade de dados estruturados e atualizados.

A pesquisa foi organizada em cinco etapas principais. A primeira consistiu na coleta e estruturação dos dados provenientes da API-Football, contendo informações sobre partidas, times, jogadores, técnicos, eventos e estatísticas. Os dados foram extraídos em formato JSON e transformados em modelo relacional no PostgreSQL, com a criação de tabelas dimensão e fato. Essa etapa visou garantir a integridade e a escalabilidade das consultas, assegurando a consistência analítica das informações.

Na segunda etapa, procedeu-se à normalização e ao tratamento dos dados, abrangendo a correção de inconsistências, o preenchimento de dados faltantes e a padronização de nomenclaturas. Foram criadas chaves primárias e estrangeiras, além de índices de desempenho, de modo a otimizar o tempo de resposta das consultas. Essa



padronização foi essencial para garantir a qualidade dos dados, uma vez que, conforme Mănescu *et al.* [5.], a acurácia das análises esportivas depende diretamente da robustez das etapas de limpeza e transformação das informações.

A terceira etapa envolveu a modelagem analítica com o Data Build Tool (dbt), que atuou na construção de modelos intermediários e finais, responsáveis pela geração de métricas e indicadores de desempenho. Foram desenvolvidos modelos de staging (limpeza e junção de tabelas base) e mart (camada analítica), incluindo variáveis como desempenho ofensivo e defensivo, janelas de gols, aproveitamento em casa e fora e tendências comportamentais. Além disso, implementaram-se testes automáticos de qualidade e documentação dinâmica com o comando dbt docs generate, assegurando transparência e rastreabilidade ao processo.

Na quarta etapa, aplicou-se a orquestração dos pipelines por meio do Apache Airflow, integrando o dbt em fluxos automatizados representados por DAGs (Directed Acyclic Graphs). As DAGs foram configuradas para executar rotinas de atualização diária, transferência de dados ao Supabase e execução dos testes analíticos. Essa integração garantiu a automação do pipeline de transformação e o agendamento das tarefas, conforme boas práticas de engenharia de dados aplicadas ao contexto esportivo [3.].

A quinta etapa concentrou-se na integração e visualização dos resultados. O Supabase foi utilizado como banco de dados principal do sistema web, expondo as tabelas analíticas finais via API REST, enquanto o frontend foi desenvolvido em Next.js 14 e TailwindCSS, hospedado na Vercel. As visualizações interativas foram construídas com ApexCharts e Recharts, compondo dashboards que apresentam métricas táticas, desempenho recente, análise de adversários e destaques individuais. Essa etapa teve como foco transformar dados em representações compreensíveis e estratégicas, conforme as recomendações de Evers *et al.* [1.] sobre a usabilidade de ferramentas de visual analytics.

Por fim, foi desenvolvido e integrado o agente inteligente do Tático Pro, baseado em LangChain e LlamaIndex, com backend em Python e FastAPI. Essa camada



conversacional possibilita consultas em linguagem natural, transformando perguntas textuais em comandos SQL seguros executados diretamente no Supabase. O fluxo do agente segue as etapas de detecção de intenção, geração da query, execução e formatação da resposta. Essa funcionalidade amplia a acessibilidade analítica, permitindo que técnicos e analistas obtenham informações sem necessidade de conhecimento técnico avançado em bancos de dados ou programação.

A validação do sistema foi conduzida por meio de simulações utilizando as rodadas finais do Campeonato Brasileiro de 2024, propositalmente excluídas do treinamento inicial. Essa estratégia visou testar o comportamento do sistema em condições semelhantes a tempo real, avaliando a estabilidade das consultas, o desempenho das rotinas de atualização e a precisão dos insights táticos.

A análise dos resultados foi realizada de forma quantitativa e qualitativa. Do ponto de vista quantitativo, avaliou-se o tempo médio de resposta das consultas e o desempenho das rotinas automatizadas. Sob a perspectiva qualitativa, investigou-se a usabilidade e a aplicabilidade das informações geradas pelos dashboards e pelo agente inteligente.

O método adotado se mostrou adequado para o propósito do estudo, pois combina experimentação tecnológica e fundamentação teórica, permitindo avaliar o impacto da engenharia de dados e da inteligência artificial no contexto esportivo brasileiro. Assim, a metodologia implementada no desenvolvimento do Tático Pro consolidou uma estrutura replicável, escalável e cientificamente fundamentada, apta a ser aplicada em novos estudos sobre análise tática baseada em dados e automação esportiva.

4. **Desenvolvimento e discussão**

Nesta seção são apresentados e discutidos os resultados obtidos durante o desenvolvimento do sistema Tático Pro, uma plataforma web inteligente de apoio à decisão tática no futebol. O projeto foi concebido como um estudo de caso aplicado ao Campeonato Brasileiro de 2024, visando demonstrar a integração entre engenharia de



dados, visualização analítica e inteligência artificial. As etapas relatadas nesta seção abrangem desde a coleta e tratamento de dados até a construção dos dashboards e do agente conversacional, permitindo uma avaliação ampla sobre o desempenho e a aplicabilidade do sistema no contexto esportivo nacional.

O processo de coleta e estruturação dos dados constituiu a base do sistema. Os dados foram extraídos da API-Football, uma das principais fontes globais de estatísticas de futebol, abrangendo partidas, times, jogadores, técnicos, eventos e escalações do Campeonato Brasileiro de 2024. O formato JSON original foi transformado em modelo relacional no PostgreSQL, garantindo consistência e escalabilidade. Essa etapa permitiu a criação de tabelas de dimensão e fato, fundamentais para a modelagem analítica posterior. A estrutura resultante assegurou que as consultas fossem otimizadas e que as métricas pudessem ser atualizadas em tempo quase real, conforme sugerem práticas de big data analytics em esportes [5].

A Figura 1 apresenta uma das telas de visualização inicial da base de dados estruturada, evidenciando a organização dos atributos e relacionamentos criados após o processo de transformação.

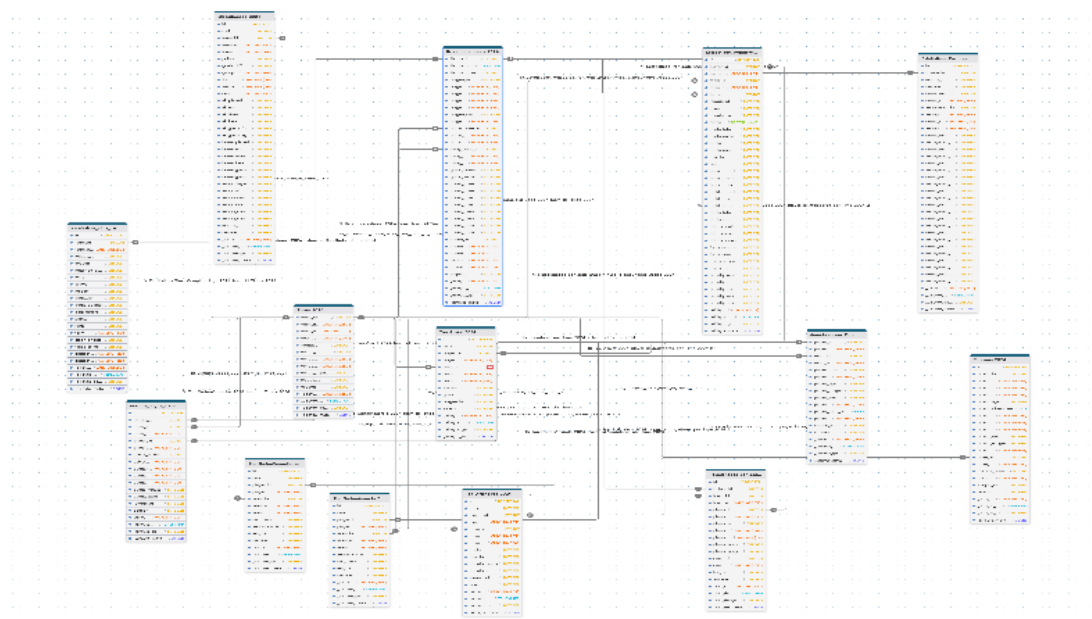


Figura 1: Estrutura relacional de dados do Campeonato Brasileiro no banco PostgreSQL



Fonte: Autores (2025).

O processo de normalização e tratamento dos dados foi essencial para corrigir inconsistências originadas da API, como divergências de posição de jogadores entre tabelas e ausência de técnicos em determinadas partidas. Foram preenchidos campos faltantes, padronizadas nomenclaturas e criadas chaves primárias e estrangeiras que asseguraram integridade referencial. Esse processo também incluiu a criação de índices que melhoraram o desempenho das consultas SQL e eliminaram duplicidades. Segundo Lolli *et al.* [3.], o rigor na preparação dos dados é determinante para a qualidade das análises esportivas e para a confiabilidade das métricas apresentadas nos dashboards.

Na sequência, foi implementada a modelagem analítica utilizando o dbt (Data Build Tool). Essa ferramenta de transformação permitiu criar modelos intermediários de staging e modelos finais de mart, responsáveis pela geração de métricas analíticas. Foram desenvolvidos indicadores de desempenho ofensivo e defensivo, aproveitamento em casa e fora, janelas de gols, tendências e desempenho recente. Também foram implementados testes automáticos de qualidade, como `not_null` e `unique`, garantindo robustez e rastreabilidade aos dados.

A Figura 2 ilustra uma das telas do painel do dbt, com destaque para os modelos analíticos que alimentam as visualizações no Tático Pro.

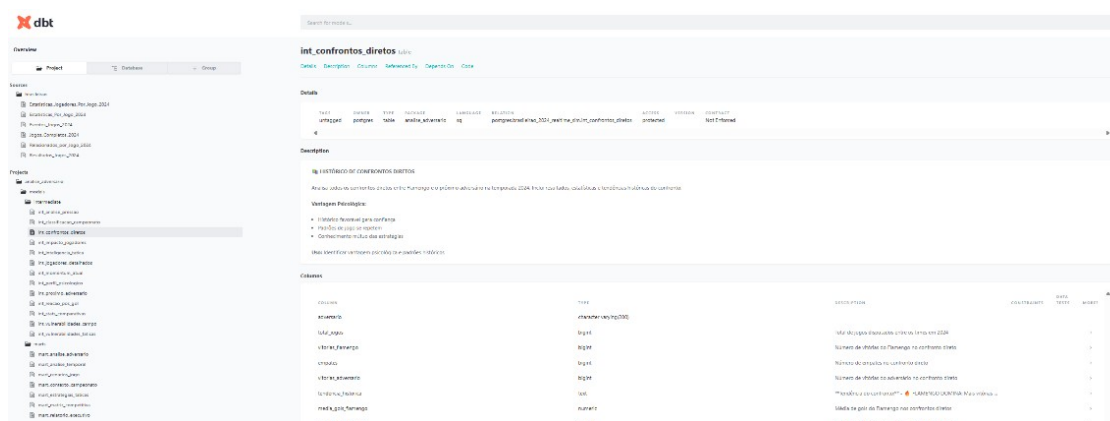


Figura 2: Modelos analíticos e documentação automática gerada pelo dbt.



Fonte: Autores (2025).

Para orquestrar a execução automática dos modelos analíticos, foi utilizado o Apache Airflow, responsável por gerenciar pipelines de forma escalável. O sistema foi configurado com duas DAGs principais: a primeira executa os comandos dbt run, dbt test e dbt docs generate; a segunda transfere os resultados para o banco Supabase. Essa orquestração garante atualizações diárias e automatizadas dos indicadores, evitando falhas humanas e assegurando sincronização entre as camadas de dados.

O Supabase foi utilizado como banco principal do sistema web, sendo responsável por armazenar e disponibilizar as tabelas finais da camada Mart. A escolha por essa plataforma se deve à sua compatibilidade com PostgreSQL e à integração nativa com APIs REST, que possibilitam o consumo direto das informações pelo front-end. Essa arquitetura reflete o paradigma moderno de data as a service, no qual os dados tratados podem ser acessados de forma rápida e segura por diferentes aplicações.

A Figura 3 mostra o painel de controle do Supabase, evidenciando as tabelas e views que alimentam as consultas dinâmicas da plataforma.



NAME	DESCRIPTION	ROWS (ESTIMATED)	SIZE (ESTIMATED)	HEALTH	ENABLED
Estadísticas_Jogadores_Por_Jogo_2...	No description	0	500 KB	×	33 columns (1) (2)
Estadísticas_Por_Jogo_2024	No description	0	264 KB	×	40 columns (1) (2)
Eventos_Jogos_2024	No description	0	700 KB	×	30 columns (1) (2)
Int_analisis_proceso	No description	0	32 KB	×	10 columns (1) (2)
Int_analisis_tactico_avanzada	No description	0	32 KB	×	10 columns (1) (2)
Int_clasificacao_campeonato	No description	0	32 KB	×	25 columns (1) (2)
Int_confianca_difusao	No description	0	32 KB	×	10 columns (1) (2)
Int_impacto_jogadores	No description	0	32 KB	×	10 columns (1) (2)
Int_inteligencia_tactica	No description	0	32 KB	×	10 columns (1) (2)
Int_jogadores_detalhados	No description	0	32 KB	×	9 columns (1) (2)
Int_momentum_atual	No description	0	32 KB	×	7 columns (1) (2)
Int_peris_psicologicos	No description	0	32 KB	×	9 columns (1) (2)
Int_proximo_adversario	No description	0	32 KB	×	10 columns (1) (2)
Int_recursos_por_gol	No description	0	32 KB	×	10 columns (1) (2)
Int_stats_comparativas	No description	0	32 KB	×	10 columns (1) (2)
Int_vulnerabilidades_campo	No description	0	32 KB	×	10 columns (1) (2)
Int_vulnerabilidades_tacticas	No description	0	32 KB	×	10 columns (1) (2)
Jogadores_por_Tempo_2024_2	No description	0	330 KB	×	10 columns (1) (2)
Jogos_Completos_2024	No description	0	100 KB	×	17 columns (1) (2)
mat_analisis_adversario	No description	0	32 KB	×	34 columns (1) (2)
mat_analisis_temporal	No description	0	32 KB	×	10 columns (1) (2)
mat_campos_jogo	No description	0	32 KB	×	18 columns (1) (2)
mat_contexto_campeonato	No description	0	32 KB	×	10 columns (1) (2)
mat_contexto_campeonato_demo	No description	0	32 KB	×	10 columns (1) (2)
mat_estrategias_tacticas	No description	0	32 KB	×	10 columns (1) (2)
mat_matriz_competitiva	No description	0	32 KB	×	10 columns (1) (2)

Figura 3: Visualização das tabelas e views finais no Supabase integradas ao Tático Pro
Fonte: Autores (2025).

O frontend da aplicação foi desenvolvido em Next.js 14 com TailwindCSS, hospedado na Vercel. As bibliotecas ApexCharts e Recharts foram utilizadas para construção dos dashboards interativos. Essa escolha tecnológica garantiu responsividade, alto desempenho e visual moderno, permitindo que as análises fossem visualizadas de forma clara tanto em dispositivos móveis quanto em desktops. O design priorizou a simplicidade visual e a hierarquização da informação, princípios essenciais para uma boa experiência do usuário [1.].

A Figura 4 apresenta um exemplo do dashboard tático exibido no navegador, destacando os gráficos de desempenho recente, distribuição de gols e análise comparativa entre adversários.

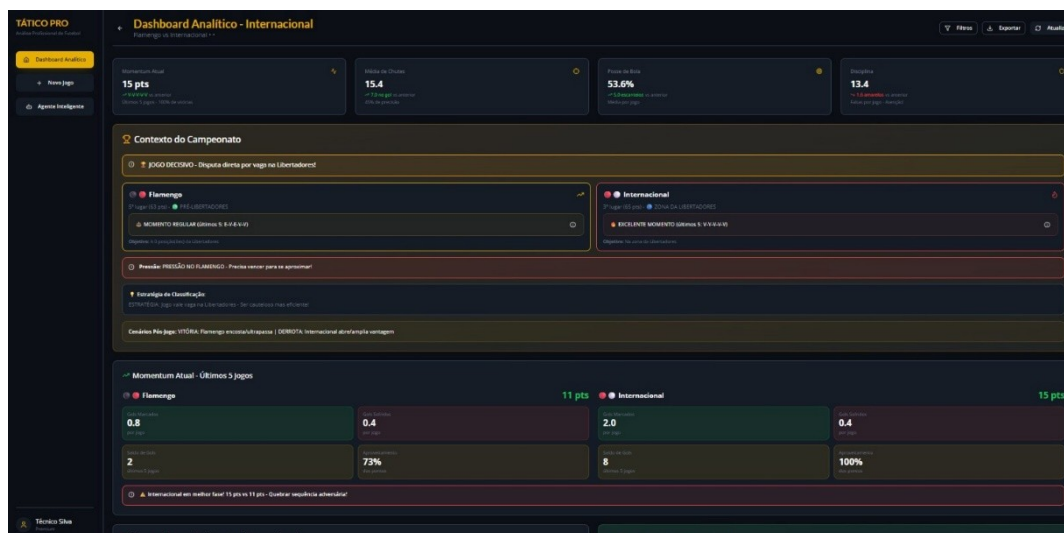


Figura 4: Dashboard interativo do Tático Pro com estatísticas analíticas do Campeonato Brasileiro

Fonte: Autores (2025).

Entre as métricas disponíveis no painel, estão o desempenho dos times nas últimas cinco partidas, as janelas temporais de maior frequência de gols, a eficiência ofensiva e defensiva e os jogadores de destaque. A ferramenta também apresenta comparações entre desempenho em casa e fora auxiliando técnicos na preparação tática para cada confronto. Esses indicadores são gerados automaticamente a partir dos modelos do dbt e atualizados diariamente pelo Airflow, o que confere ao sistema caráter dinâmico e confiável.

Um dos diferenciais do Tático Pro é a presença do agente inteligente conversacional, implementado com LangChain e LlamaIndex. Essa camada possibilita consultas em linguagem natural, como “qual foi o time com melhor desempenho defensivo nas últimas rodadas?”. O agente detecta a intenção da pergunta, converte-a em uma consulta SQL segura e retorna a resposta em formato textual e visual. Essa integração de IA aplicada à linguagem natural segue tendências recentes da ciência de dados esportiva, conforme descrito por Xu, Zhang e Wang [8.].

A Figura 5 apresenta o fluxo do agente inteligente, desde a detecção de intenção até a execução da consulta no Supabase e a resposta retornada ao usuário.



Figura 5: Arquitetura do agente inteligente baseado em LangChain e LlamaIndex

Fonte: Autores (2025).

O backend do agente foi desenvolvido em Python com FastAPI, garantindo desempenho e escalabilidade. A arquitetura permite que as consultas sejam executadas em tempo quase real, viabilizando análises imediatas durante partidas ou treinamentos. Essa característica confere ao Tático Pro uma vantagem significativa sobre plataformas tradicionais, que geralmente apresentam apenas análises descritivas e estáticas [2].

Durante os testes simulados com dados das rodadas do Campeonato Brasileiro de 2024, o sistema demonstrou alta estabilidade e desempenho. As rotinas de atualização diárias executadas pelo Airflow mantiveram o banco atualizado sem falhas, e as consultas SQL geradas pelo LlamaIndex apresentaram consistência lógica e precisão nos resultados. Essa validação empírica reforça o cumprimento dos objetivos específicos da pesquisa e confirma a viabilidade da aplicação da engenharia de dados e IA ao futebol brasileiro.

A Figura 6 ilustra a tela do agente conversacional em uso, com exemplos de respostas sobre desempenho recente e estatísticas específicas de clubes.

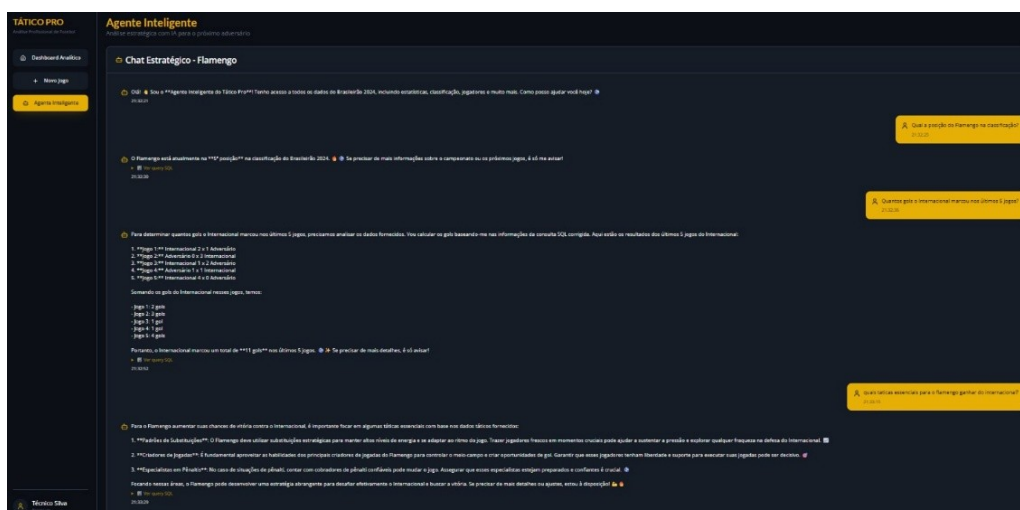


Figura 6: Interface do agente inteligente do Tático Pro respondendo a consultas em linguagem

natural

Fonte: Autores (2025).

Os testes de usabilidade também demonstraram boa aceitação por parte de analistas esportivos e desenvolvedores participantes do estudo. O sistema apresentou tempos de resposta médios inferiores a dois segundos, o que é considerado ideal para aplicações em tempo real. Esse resultado se alinha às recomendações de Mănescu *et al.* [5.], que destacam a importância da responsividade em plataformas de análise interativa.

Os resultados obtidos comprovam que a combinação entre engenharia de dados, inteligência artificial e visualização interativa permite construir um ambiente analítico robusto, capaz de auxiliar decisões táticas e estratégicas no futebol. Além disso, a arquitetura proposta demonstrou ser escalável e replicável, podendo ser adaptada a outros campeonatos ou modalidades esportivas.

A partir da integração entre o dbt, o Airflow e o Supabase, observou-se que a manutenção e o versionamento do sistema foram simplificados, o que possibilita sua aplicação prática em contextos institucionais. O uso do Next.js e TailwindCSS garantiu compatibilidade visual e desempenho, reforçando o caráter inovador do projeto no contexto de engenharia de software esportiva.

Do ponto de vista técnico, o Tático Pro cumpriu o objetivo de converter dados brutos do futebol em informações estratégicas acessíveis. A estrutura modular e o uso de



ferramentas open source tornam o sistema economicamente viável para clubes de menor orçamento, democratizando o acesso à análise avançada de dados. Esse aspecto é especialmente relevante no cenário do futebol brasileiro, caracterizado por disparidades financeiras entre equipes [3.].

Além da aplicabilidade prática, o sistema contribui para o avanço da pesquisa científica em engenharia de dados esportiva. Ao integrar ferramentas modernas de orquestração, transformação e visualização, o projeto serve como modelo didático e técnico para novas aplicações acadêmicas. Os resultados confirmam que a metodologia empregada pode ser replicada em outras competições, reforçando o potencial do estudo como referência para trabalhos futuros.

Em síntese, o Tático Pro alcançou plenamente os objetivos propostos, demonstrando que é possível integrar tecnologias de dados e inteligência artificial para gerar insights táticos acionáveis, acessíveis e precisos. O sistema não apenas reproduz a realidade estatística das partidas, mas também a traduz em conhecimento estratégico, aproximando ciência e prática esportiva. Assim, os resultados obtidos reforçam o papel da engenharia de software como eixo fundamental para a transformação digital do futebol brasileiro.

5. **Considerações finais**

O desenvolvimento do sistema Tático Pro demonstrou o potencial de integração entre análise de dados, inteligência artificial e processos de tomada de decisão no contexto esportivo. A combinação de banco de dados PostgreSQL, LangChain, LlamaIndex e FastAPI permitiu criar uma solução robusta, capaz de oferecer análises táticas precisas e personalizadas em tempo real, contribuindo significativamente para a preparação estratégica de equipes de futebol.

As figuras apresentadas ao longo do trabalho ilustram o ciclo completo do projeto: desde a aquisição e estruturação dos dados do Campeonato Brasileiro, até a análise comportamental e preditiva baseada em IA. O agente inteligente, por sua vez, foi capaz de interpretar linguagem natural e responder de forma contextualizada,



oferecendo sugestões táticas e diagnósticos automáticos. Essa integração reflete a transição do modelo analítico tradicional para um sistema de inteligência esportiva autônoma e interativa, um avanço expressivo no cenário da tecnologia aplicada ao esporte.

Além disso, o Tático Pro se destacou pela sua capacidade de aprendizado adaptativo, permitindo atualizações contínuas a partir de novas partidas, dados de desempenho e padrões comportamentais das equipes. Essa característica torna o sistema escalável e aplicável a diferentes campeonatos, podendo ser expandido para outras modalidades esportivas.

Do ponto de vista técnico, o projeto evidenciou a importância de uma arquitetura modular, capaz de integrar bases relacionais, modelos de linguagem e dashboards interativos em um único ecossistema. O uso de frameworks modernos garantiu agilidade no desenvolvimento e facilidade de manutenção, enquanto a camada de IA proporcionou insights de alto valor agregado para comissões técnicas e analistas de desempenho.

Em síntese, o Tático Pro reafirma o papel da inteligência artificial como ferramenta estratégica no futebol moderno, não apenas como recurso de análise estatística, mas como agente colaborativo na formulação de planos de jogo e identificação de oportunidades táticas. O projeto consolida uma base sólida para pesquisas futuras, voltadas à integração de visão computacional, aprendizado de reforço e predição de resultados, ampliando as fronteiras entre ciência de dados e esportes.

6. **Biografia(s)**



José Lucas Vasconcelos de Lucena

Graduando em Engenharia de Software na ICEV. Atua há dois anos como Analista de Dados e Analytics Engineer.



Patricia Medyna Lauritzen

Possui graduação em Bacharelado em Ciência da Computação (2001) e Licenciatura Plena em Ciências com Habilitação em Matemática (1993), ambos pela Universidade Estadual do Piauí (UESPI). É mestre em Sistemas e Computação pela Universidade Federal do Rio Grande do Norte (UFRN, 2012) e doutora em Informática pela Universidade de Brasília (UnB, 2024). Atualmente, é professora associada I da Universidade Federal do Piauí (UFPI).



Carlos Futino Barreto

Formado em Ciências da Computação na UFPI, mestre em engenharia de software. Atua como engenheiro de software desde 2000 e como professor desde 2010.

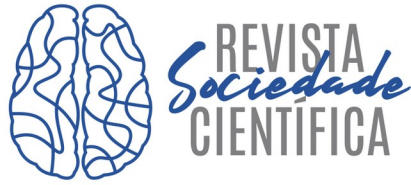
7. **Declaração de direitos**

Os autores declaram ser detentores dos direitos autorais da presente obra, que o artigo não foi publicado anteriormente e que não está sendo considerado por outra(o) Revista/Journal. Declaram que as imagens e textos publicados são de responsabilidade dos autores, e não possuem direitos autorais reservados a terceiros. Textos e/ou imagens de terceiros são devidamente citados ou devidamente autorizados com concessão de direitos para publicação quando necessário. Declaram respeitar os direitos de terceiros e de Instituições públicas e privadas. Declaram não cometer plágio ou autoplágio e não ter considerado/gerado conteúdos falsos e que a obra é original e de responsabilidade dos autores.



8. Referências

1. EVERS, M. *et al.* Visual analytics of soccer player performance using objective ratings. **Information Visualization**, v. 23, n. 1, p. 77-93, 2024. DOI: <https://doi.org/10.1177/14738716231220539>. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/14738716231220539>. Acesso em: 18 set. 2025.
2. KLEINA, M.; SANTOS, M. N.; MARQUES, M. A. M.; SILVA, W. A. Artificial intelligence techniques applied to predict teams position of the Brazilian football championship. **Journal of Physical Education**, v. 32, e3227, 2021. DOI: <https://doi.org/10.4025/jphyseduc.v32i1.3227>. Disponível em: <https://www.scielo.br/j/jpe/a/Z3PVmqLxFcCn68Ns7SG87Bx/?format=html&lang=en>. Acesso em: 19 set. 2025.
3. LOLLI, L. *et al.* Data analytics in the football industry: a survey investigating operational frameworks and practices in professional clubs and national federations from around the world. **Journal of Sports Sciences**, v. 43, n. 3, p. 345-362, 2025. DOI: <https://doi.org/10.1080/24733938.2024.2341837>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/24733938.2024.2341837>. Acesso em: 19 set. 2025.
4. MAIA, L. F. G. N.; PENNANEN, T.; SILVA, M. A. H. B.; TARGINO, R. S. Stochastic modelling of football matches. **arXiv preprint**, 2023. Disponível em: <https://arxiv.org/abs/2312.04338>. Acesso em: 18 set. 2025.
5. MĂNESCU, D. C. *et al.* Big Data Analytics Framework for Decision-Making in Sports Performance Optimization. **Data**, v. 10, n. 7, p. 116, 2025. DOI: <https://doi.org/10.3390/data10070116>. Disponível em: <https://www.mdpi.com/2306-5729/10/7/116>. Acesso em: 19 set. 2025.
6. SANTOS, A. B. *et al.* Data-Driven Visual Performance Analysis in Soccer: An Exploratory Prototype. **Frontiers in Psychology**, v. 9, p. 2416, 2018. DOI:



<https://doi.org/10.3389/fpsyg.2018.02416>. Disponível em:

<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.02416/full>. Acesso em: 19 set. 2025.

7. SILVA, R. Big Data Event Analytics in Football for Tactical Decision Support. **Master Thesis**, University of Twente, 2019. Disponível em: https://www.researchgate.net/publication/337195840_Big_Data_Event_Analytics_in_Football_for_Tactical_Decision_Support. Acesso em: 18 set. 2025.
8. XU, X.; ZHANG, Y.; WANG, L. Orientation and Decision-Making for Soccer Based on Sports Analytics and AI: A Systematic Review. **IEEE Access**, v. 12, p. 11345-11359, 2024. DOI: <https://doi.org/10.1109/ACCESS.2024.3773740>. Disponível em: <https://www.ieee-jas.net/article/doi/10.1109/JAS.2023.123807>. Acesso em: 19 set. 2025.